



Article

A Hierarchical Information Extraction Method for Large-Scale Centralized Photovoltaic Power Plants Based on Multi-Source Remote Sensing Images

Fan Ge ^{1,2}, Guizhou Wang ^{1,3}, Guojin He ^{1,2,3,*}, Dengji Zhou ^{1,2} , Ranyu Yin ¹ and Lianzi Tong ^{1,2}

¹ Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China

² University of Chinese Academy of Sciences, Beijing 100049, China

³ Key Laboratory of Earth Observation of Hainan Province, Hainan Research Institute, Aerospace Information Research Institute, Chinese Academy of Sciences, Sanya 572029, China

* Correspondence: hegj@aircas.ac.cn; Tel.: +86-010-82178190

Abstract: In the context of global sustainable development, solar energy is very widely used. The installed capacity of photovoltaic panels in countries around the world, especially in China, is increasing steadily and rapidly. In order to obtain accurate information about photovoltaic panels and provide data support for the macro-control of the photovoltaic industry, this paper proposed a hierarchical information extraction method, including positioning information and shape information, and carried out photovoltaic panel distribution mapping. This method is suitable for large-scale centralized photovoltaic power plants based on multi-source satellite remote sensing images. This experiment takes the three northwest provinces of China as the research area. First, a deep learning scene classification model, the EfficientNet-B5 model, is used to locate the photovoltaic power plants on 16-m spatial resolution images. This step obtains the area that contains or may contain photovoltaic panels, greatly reducing the study area with an accuracy of 99.97%. Second, a deep learning semantic segmentation model, the U²-Net model, is used to precisely locate photovoltaic panels on 2-m spatial resolution images. This step achieves the exact extraction results of the photovoltaic panels from the area obtained in the previous step, with an accuracy of 97.686%. This paper verifies the superiority of a hierarchical information extraction method in terms of accuracy and efficiency through comparative experiments with DeepLabV3+, U-Net, SegNet, and FCN8s. This meaningful work identified 180 centralized photovoltaic power plants in the study area. Additionally, this method makes full use of the characteristics of different remote sensing data sources. This method can be applied to the rapid extraction of global photovoltaic panels.

Keywords: hierarchical information extraction; centralized photovoltaic power plants; large-scale; multi-source remote sensing images



Citation: Ge, F.; Wang, G.; He, G.; Zhou, D.; Yin, R.; Tong, L. A Hierarchical Information Extraction Method for Large-Scale Centralized Photovoltaic Power Plants Based on Multi-Source Remote Sensing Images. *Remote Sens.* **2022**, *14*, 4211. <https://doi.org/10.3390/rs14174211>

Academic Editor: Amin Beiranvand Pour

Received: 5 July 2022

Accepted: 24 August 2022

Published: 26 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Against the international background of global sustainable development and countries committed to building a community with a shared future for humankind, solar energy has unique advantages, such as inexhaustibility, no transportation, and no pollution [1]. As an important energy source to alleviate the future world energy crisis, solar energy has become a major research interest, and the photovoltaic power generation industry is in full swing. Under the guidance of various countries' photovoltaic policies [2–7] and the promotion of the market, the global photovoltaic installed capacity has grown steadily and rapidly. By the end of 2020, the global photovoltaic installed capacity exceeded 760.4 GW [8]. By the end of 2021, China's photovoltaic grid-connected installed capacity reached 306 GW [9]. How to supervise the large-scale and fast-growing photovoltaic power plants from a macro perspective is an urgent problem to be solved in the fields of land survey and resource and environmental monitoring [10]. Satellite remote sensing images have the advantages of

a large detection range, fast data acquisition, high update frequency, and freedom from ground conditions [11]. It is worth mentioning that photovoltaic power plants have an obvious texture, geometry, spectrum, and other characteristics in remote sensing images. Therefore, it is of great significance to propose a hierarchical information extraction method for centralized photovoltaic power plants based on multi-source satellite remote sensing images. This section mainly describes the research status from the aspects of traditional photovoltaic power plant identification methods, photovoltaic power plant identification methods based on deep learning, and small-target identification methods [12].

This paragraph describes traditional methods used for photovoltaic panels extraction. Zhang et al. extracted photovoltaic power plants from a Landsat-8 image using a random forest model on the Google Earth Engine platform [13]. The input of the model is the texture features calculated by the grayscale co-occurrence matrix, the normalized difference vegetation index (NDVI), the normalized difference building index (NDBI), the modified normalized difference water index (MNDWI), the reflectivity, and the thermal spectrum features. Experiments verified that the texture features of the gray-scale co-occurrence matrix can significantly improve the accuracy of the model to identify photovoltaics. However, the gray-scale co-occurrence matrix is complexly calculated pixel by pixel, based on the neighborhood around the pixel, and the calculation amount is very large, which will greatly reduce the efficiency of the model and increase the time cost. It is not suitable for large-scale extraction of photovoltaic panels. Li et al. used the spectral-information-rich Landsat-8 Operational Land Imager (OLI) data to calculate the normalized building index and brightness, two spectral characteristics, to separate the solar cell array, construction land, road, and fallow land. For high-resolution remote sensing images such as the GaoFen-1 image, the authors distinguish solar photovoltaic arrays from other ground objects based on geometric and texture features [14]. Wang et al. used a one-class support vector machine (OC-SVM) method to extract photovoltaic power stations along the coast of Jiangsu Province and used the modified optimum index factor (MOIF) for extracting photovoltaic power stations. Compared with the traditional optimal index factor (OIF) band combination method, the MOIF improves the extraction accuracy of photovoltaic panels [15]. Therefore, it is not an optimal method to extract photovoltaic panels from GaoFen satellite images that are not so rich in spectral information, relying on geometric and texture features to identify photovoltaic panels in a large range.

This paragraph describes deep learning methods used for photovoltaic panel extraction. Costa et al. used deep learning methods to select 24 study areas in Brazil to identify and monitor photovoltaic power plants in Sentinel-2 images [16]. The research compares the effects of four common frameworks, U-Net [17], DeepLabv3+ [18], Feature pyramid networks (FPN) [19], and Pyramid Scene Parsing Network (PSPNet) [20], in photovoltaic panel recognition. Although the U-Net architecture presented the best results in the photovoltaic identification process, it did not show much difference from other architectures. Therefore, when using deep learning methods for ground object information recognition, although the recognition accuracy is closely related to the structure of the model, the recognition effect is more dependent on the construction of the dataset. Using deep learning methods [21] to directly perform semantic segmentation of photovoltaic panels on high-resolution remote sensing images is not an optimal method as it may bring unnecessary computational consumption when photovoltaic panel pixels are much lower in number than non-photovoltaic panels.

This paragraph describes scene-oriented photovoltaic panel positioning methods. Wang et al. selected seven pre-trained convolutional neural network models for the scene classification of photovoltaic power plants in Landsat 8 OLI images [10]. The sample set contains 120 positive and negative sample images of 56×56 pixels in RGB bands, including as many photovoltaic panels of different forms as possible. After adopting the strategies of transfer learning and model fine-tuning, the scene recognition of photovoltaic panels was carried out. The results show that it is feasible to use the convolutional neural network model to perform scene recognition of photovoltaic panels on medium-resolution remote sensing images. How-

ever, scene classification cannot obtain fine photovoltaic boundary information, so it is far from sufficient to only perform scene recognition on photovoltaic panels.

Therefore, this paper proposes a hierarchical information extraction method of centralized photovoltaic power plants based on multi-source remote sensing images and conducts large-scale photovoltaic power plant distribution mapping [22]. The method firstly performs scene classification of photovoltaic panels in medium-resolution remote sensing image, and obtains the area containing photovoltaic panels or suspected photovoltaic panels, which greatly reduces the background area without photovoltaic panels and balances the number of positive and negative targets. We use this as a mask to accurately identify photovoltaic panels in high-resolution remote sensing images. This method not only avoids the problem of unbalanced positive and negative targets faced by using deep learning methods to directly perform semantic segmentation on photovoltaic panels, but also improves the speed of photovoltaic panel identification on a large scale, greatly improves recognition efficiency and saves computing resources.

2. Materials

2.1. Study Area

In this experiment, three provinces in northwest China were selected as the study area, as shown in Figure 1, namely, Xinjiang Uygur Autonomous Region, Qinghai Province, and Gansu Province. The three provinces account for about 30% of the Chinese land area [23]. The study area is vast and has sufficient sunlight, making it an ideal location for large-scale photovoltaic power plants. In recent years, the study area has vigorously developed the photovoltaic industry. By the end of 2021, the cumulative installed capacity of photovoltaics in the three provinces had far exceeded 10 million kilowatts [24]. However, the typical terrain of Xinjiang Uygur Autonomous Region is mostly desert, and the typical terrain of Qinghai Province and Gansu Province is mostly alpine and Gobi. It is difficult to count the construction of centralized photovoltaic power plants in a timely manner using manpower. Therefore, three provinces in northwestern China were selected as the study area in this experiment.

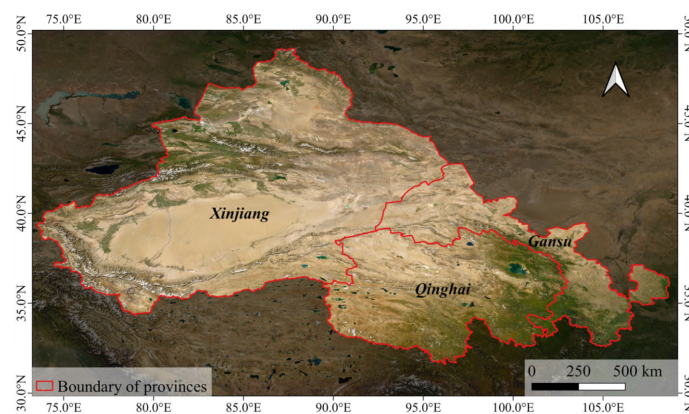


Figure 1. Geographical location of the study area.

2.2. Data

Chinese satellite images were used in this study, including 16-m medium-resolution images and 2-m high-resolution images. The medium-resolution remote sensing images used in this experiment were GaoFen-6 Wide Field Vision (WFV) images. The high-resolution remote sensing images used in this experiment were GaoFen-6 Panchromatic Multi-Spectral (PMS) and GaoFen-1/1B/1C/1D PMS images.

The GaoFen-6 satellite is equipped with a 2-m panchromatic/8-m multispectral high-resolution camera that can shoot PMS images and a 16-m multispectral medium-resolution camera that can shoot WFV images. GaoFen-6 PMS image contains 4 bands with a spatial resolution of 2 m; the WFV image of GaoFen-6 contains 8 bands with a spatial resolution of 16 m. The spectral parameters of GaoFen-6 are shown in Table 1.

Table 1. GaoFen-6 spectral parameters.

GaoFen-6 PMS			GaoFen-6 WFV		
Spectral range	PAN	0.45–0.90 μm	Spectral range	B1	0.45–0.52 μm
		0.45–0.52 μm		B2	0.52–0.59 μm
	MS	0.52–0.59 μm		B3	0.63–0.69 μm
		0.63–0.69 μm		B4	0.77–0.89 μm
		0.76–0.89 μm		B5	0.69–0.73 μm
Spatial resolution	PAN	2 m	Spatial resolution	B6	0.73–0.77 μm
	MS	8 m		B7	0.40–0.45 μm
				B8	0.59–0.63 μm
					16 m

The GaoFen-1 satellite is equipped with two 2-m resolution panchromatic/8-m resolution multispectral cameras that can shoot PMS images and four 16-m resolution multispectral cameras that can shoot WFV images. The GaoFen-1 PMS image contains 4 bands with a spatial resolution of 2 m; the WFV image of GaoFen-1 contains 4 bands with a spatial resolution of 16 m. The spectral parameters of GaoFen-1 are shown in Table 2.

Table 2. GaoFen-1 spectral parameters.

GaoFen-1 PMS			GaoFen-1 WFV		
Spectral range	PAN	0.45–0.90 μm	Spectral range	B1	0.45–0.52 μm
		0.45–0.52 μm		B2	0.52–0.59 μm
	MS	0.52–0.59 μm		B3	0.63–0.69 μm
		0.63–0.69 μm		B4	0.77–0.89 μm
Spatial resolution	PAN	2 m	Spatial resolution		
	MS	8 m			16 m

2.3. Dataset Construction

This experiment constructs two types of datasets for scene classification and for semantic segmentation.

2.3.1. Sixteen-Meter Resolution Dataset for Scene Classification

The data source of the scene classification dataset is the GaoFen-6 WFV image with a spatial resolution of 16 m. The dataset includes two categories: one is the scene area that does not contain photovoltaic panels, and the other is scene areas that contain or are suspected to contain photovoltaic panels. The purpose of the spatial information positioning of photovoltaic panels based on medium-resolution remote sensing images is to find the locations of as many photovoltaic panels as possible. On the basis of ensuring that the photovoltaic panels are not missed, the false positives of the photovoltaic panels are reduced. Additionally, in the process of visually interpreting remote sensing images with a spatial resolution of 16 m, there are ground objects that are easily confused with photovoltaic panels, such as greenhouses. Therefore, in the process of constructing the scene classification dataset, the ground objects suspected of being photovoltaic panels are also divided into categories including photovoltaic panels. Figure 2 shows part of the scene classification dataset.

The distribution of the scene classification dataset is shown in Figure 3. Red dots represent scene samples that contain photovoltaic panels, and green dots represent scene samples that do not contain photovoltaic panels. After image flipping, image mirroring, color adjustment and other sample enhancement [25–27] operations, a total of 22,025 samples with a size of 128×128 pixels were obtained. There are 5590 samples containing photovoltaics and 16,435 samples not containing photovoltaics.

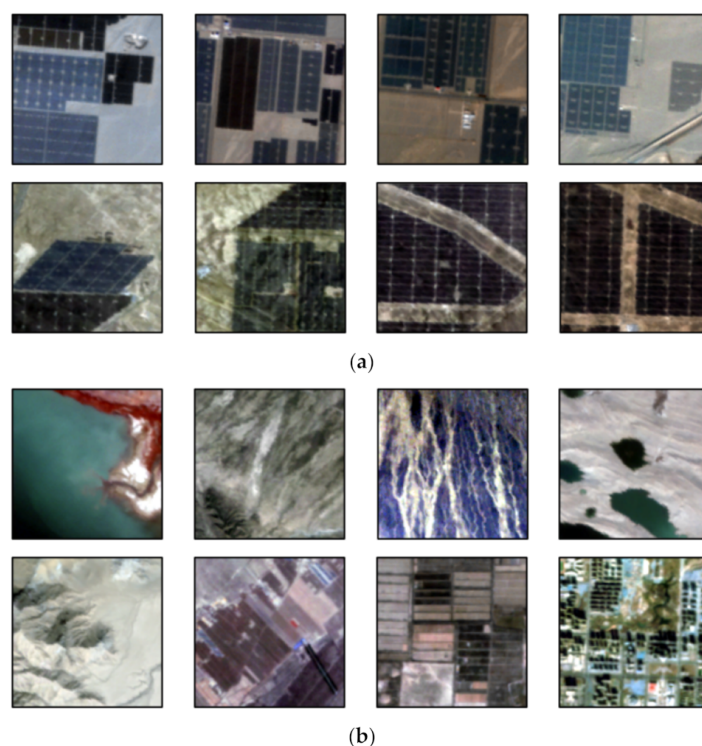


Figure 2. Scene classification dataset display: (a) PV; (b) non-PV.

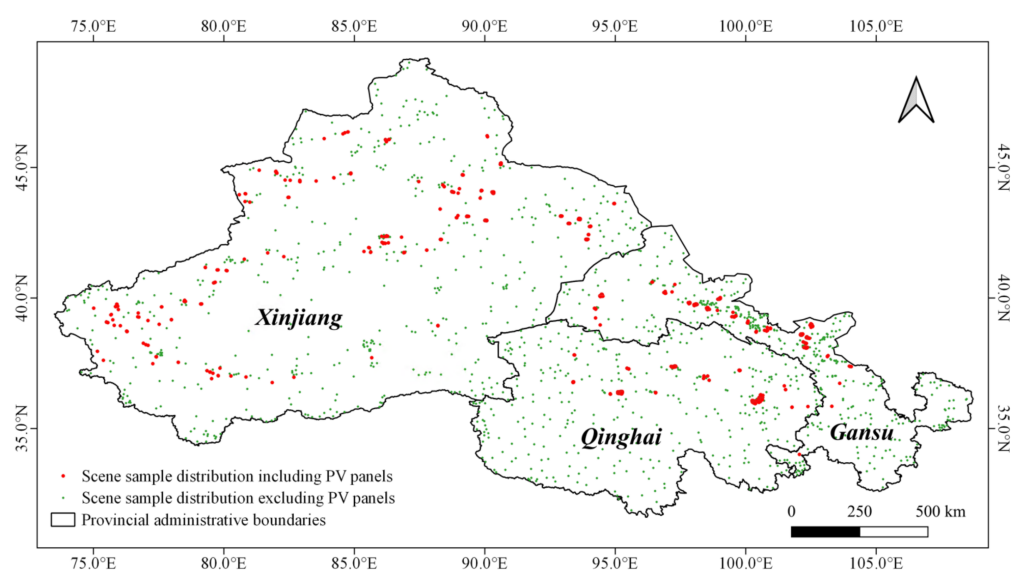


Figure 3. Sample distribution map for scene classification dataset.

2.3.2. Two-Meter Resolution Dataset for Semantic Segmentation

The data source of the semantic segmentation dataset consists of GaoFen-6 and GaoFen-1/1B/1C/1D PMS images with a spatial resolution of 2 m. The pixels in the dataset are divided into two categories: One is photovoltaic panel pixels, and the other is non-photovoltaic panel pixels. Figure 4 shows part of the semantic segmentation dataset.

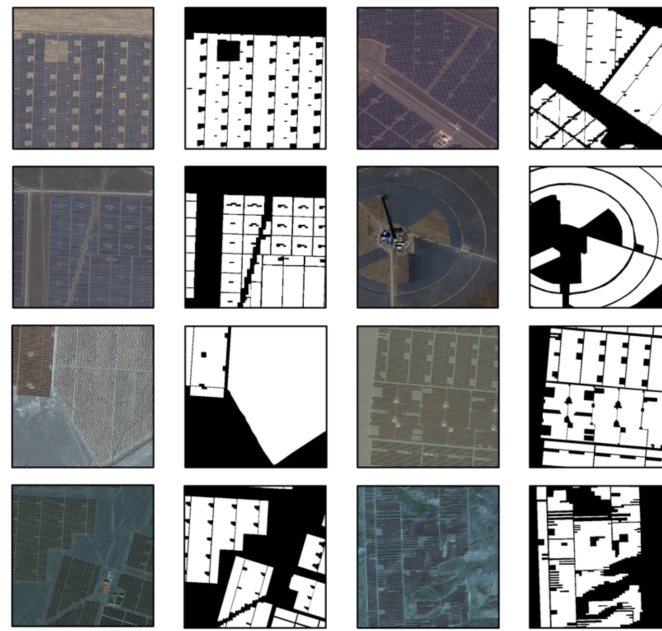


Figure 4. Semantic segmentation dataset display.

The semantic segmentation dataset contains a total of 2671 samples with a size of 512×512 pixels.

3. Methods

In order to make full use of the characteristics of remote sensing images with different spatial resolutions, improve the accuracy and efficiency of extracting photovoltaic panels, and save computing resources, this paper proposes a hierarchical information extraction method for large-scale centralized photovoltaic power plants based on multi-source remote sensing images.

This method is divided into two main parts: in the first part, the deep learning scene classification model is used to locate the photovoltaic power plants on the 16-m spatial resolution image; in the second part, the semantic segmentation model is used to extract photovoltaic panels precisely on the 2-m spatial resolution image.

The technical framework of this method is shown in Figure 5:

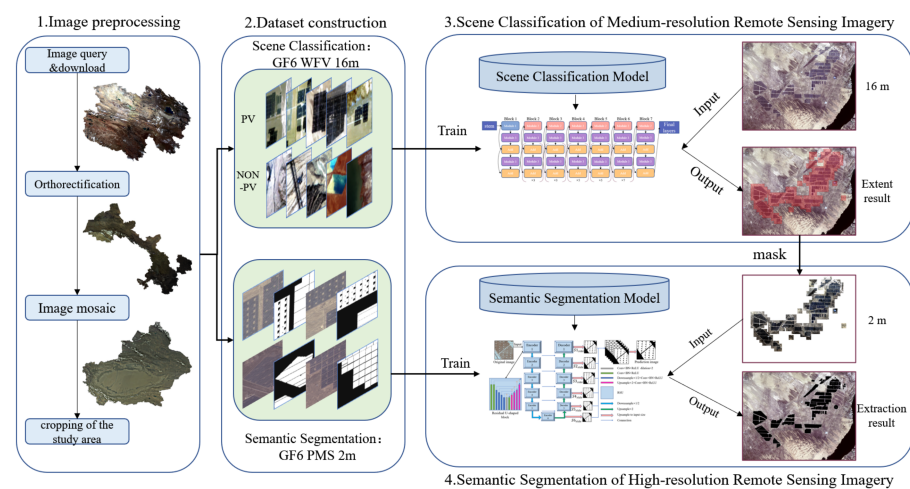


Figure 5. Framework for hierarchical information extraction of large-scale centralized photovoltaic power plants.

The first step is remote sensing image preprocessing. For remote sensing images with a spatial resolution of 16 m, the original images were subjected to orthorectification, bit-depth adjustment, image mosaicking, and cropping of the study area; for remote sensing images with a spatial resolution of 2 m, the original images were subjected to orthographic correction, sharpening fusion, bit-depth adjustment, image mosaicking, and cropping of the study area.

The second step is dataset construction, as mentioned in Section 2. Based on images with a spatial resolution of 16 m, a scene classification dataset is constructed. The dataset includes two types of areas, including photovoltaic samples and areas without photovoltaic samples; based on images with a spatial resolution of 2 m, a semantic segmentation dataset is constructed. The pixels in the dataset are divided into photovoltaic pixels and non-photovoltaic pixels.

The third step is to classify scenes with or without photovoltaic panels based on 16-m spatial resolution remote sensing images. In order to quickly locate the spatial information of the centralized photovoltaic panels and greatly reduce the scope of the study area, a scene classification model, such as the EfficientNet-b5 model, is trained based on the scene classification dataset. The model is based on a model trained on tens of millions of ImageNet datasets. The input to the model is a 16-m spatial resolution image, and the output of the model is the geographical range of the area that contains or may contain photovoltaic panels. The output geographic range is saved as a shape file through a script file written in advance.

The fourth step is to accurately identify photovoltaic panels based on 2-m spatial resolution remote sensing images. A semantic segmentation model such as the U²-Net model is trained based on the semantic segmentation dataset. According to the geographic range output in step three, we directly read the high-resolution remote sensing image of the corresponding area. The input to the model is two-meter resolution images of the area that contains or may contain photovoltaic panels, and the output of the model is the exact extraction result of the photovoltaic panels.

This framework enables fast and precise extraction of centralized photovoltaic panels in a large-scale area.

3.1. EfficientNet

EfficientNet is a new network model proposed by Tan et al. in 2019 [28]. The main innovation of EfficientNet is to improve the accuracy of the model through model scaling. The backbone network of the model uses MBConv in MobileNet V2 [29] and uses the squeeze and excitation method in SENet [30] to optimize the network structure. The model uses grid search to balance the scaling ratios of the three dimensions of network width, depth, and resolution to obtain higher accuracy and efficiency. Tan et al. obtained the B1-B7 network by extending the baseline network, and applied the series network of EfficientNet to ImageNet dataset. Comparing the series of EfficientNet networks with classic networks such as ResNet [31], DenseNet [32], etc., EfficientNet rapidly improves accuracy without a significant increase in the number of parameters. Therefore, the effect of the EfficientNet model is far better than other classical models. Considering the influencing factors such as model accuracy, efficiency, computing power, etc., the EfficientNet-B5 model with relatively high accuracy and relatively few parameters is used in this experiment to locate the spatial information of photovoltaic panels. Its network structure is shown in Figure 6.

3.2. U²-Net

The U²-Net is a simple and powerful deep network architecture for salient object detection proposed by Qin et al. in 2020 [33]. The model structure is a two-layer nested U-shaped structure, which is upgraded and improved on the basis of the U-Net network [17]. The Residual U-shaped module (RSU) included in the U²-Net network is able to capture multi-scale deep features.

The double-nested U-shaped structure of U²-Net has two significant advantages: First, the RSU module included in the model mixes receptive fields of different sizes, which can capture more and richer contextual information at different scales; Second, although the model increases the depth of the network, the entire model architecture does not significantly increase the computational cost due to the pooling operation used in the RSU module.

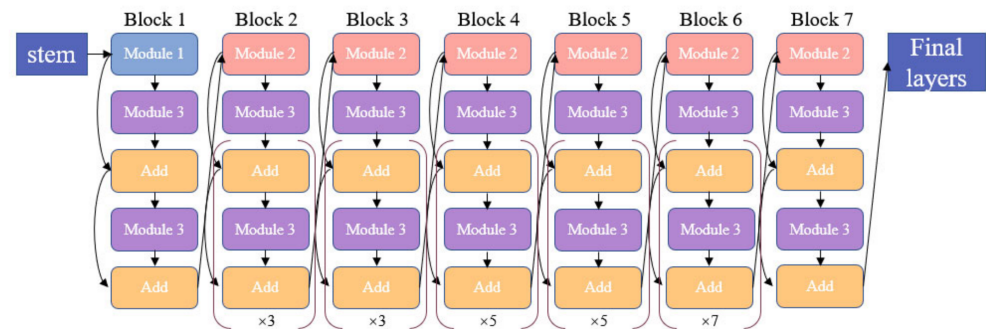


Figure 6. The network structure of EfficientNet-B5.

Considering the influencing factors such as model accuracy and calculation cost, the U²-Net model is used in this experiment to accurately identify photovoltaic panels, and its network structure is shown in Figure 7.

3.3. Accuracy Evaluation Method

The rapid identification method for large-scale centralized photovoltaic power plants proposed in this paper is divided into two steps: photovoltaic power plant spatial information positioning and photovoltaic panel accurate identification. Therefore, the accuracy evaluation method of this experiment is also divided into two parts.

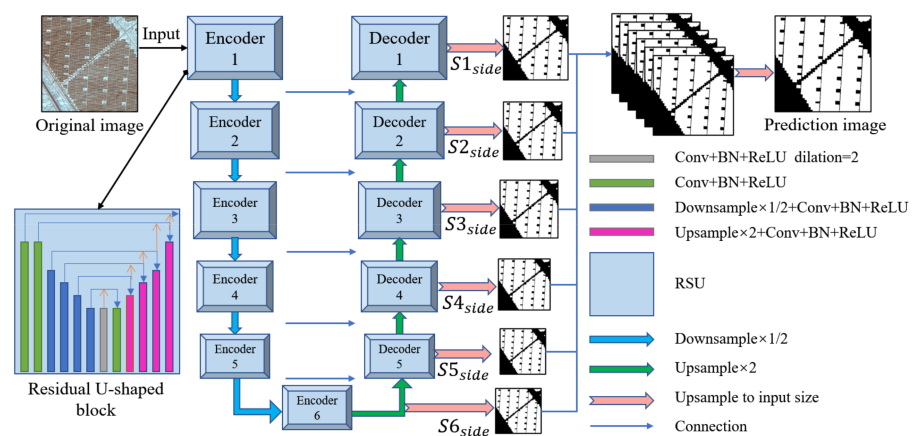


Figure 7. The network structure of U²-Net.

For the spatial information positioning process of photovoltaic power plants based on medium-resolution remote sensing images, two indicators, Precision₁ and Recall₁, are used to evaluate the accuracy. This process is concerned with whether the scenes involving photovoltaics are correctly predicted. Therefore, these two evaluation indicators are aimed at scenes including photovoltaics.

Precision₁ represents the proportion of a category that is correctly classified in the classification results.

Recall₁ represents the proportion of a category that is correctly predicted in reality.

The evaluation process selects part of the study area as the test area. The three result parameters in Table 3 were obtained by manual interpretation.

Table 3. Scene classification prediction result parameters.

Number of scenes: Reality: PV Result: PV	Number of scenes: Reality: non-PV Result: PV	Number of scenes: Reality: PV Result: non-PV
N_1	N_2	N_3

The parameter N_1 represents the number of scenes that include photovoltaics both in reality and in the predicted result; the parameter N_2 represents the number of scenes which that do not include photovoltaics in reality but include photovoltaics in the predicted result; and the parameter N_3 represents the number of scenes that include photovoltaics in reality but do not include photovoltaics in the predicted result.

$$\text{Precision}_1 = \frac{N_1}{N_1 + N_2} \times 100\% \quad (1)$$

$$\text{Recall}_1 = \frac{N_1}{N_1 + N_3} \times 100\% \quad (2)$$

The evaluation indicator Precision_1 can evaluate the prediction accuracy of the model for photovoltaic category. The evaluation index Recall_1 can measure the impact of scene classification on the extraction of photovoltaic panels.

For the accurate identification process of photovoltaic panels based on high-resolution remote sensing images, several evaluation indicators commonly used in semantic segmentation are used to evaluate the accuracy, including Accuracy, Precision_2 , Recall_2 , F1, and Intersection over Union (IoU). The above evaluation indicators are all calculated based on the confusion matrix [34], as shown in Table 4.

Table 4. Schematic diagram of confusion matrix.

Confusion Matrix		Prediction Result	
		Positive Sample	Negative Sample
Ground truth	PositivesSample	TP	FN
	Negative sample	FP	TN

Accuracy indicates the percentage of correctly predicted samples in the total samples. The value ranges from 0 to 1. The larger the value, the better the prediction effect:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (3)$$

Precision_2 indicates the proportion of all the samples predicted to be positive samples, the true label is a positive sample, the value range is 0 to 1, and the larger the value, the better the prediction effect:

$$\text{Precision}_2 = \frac{TP}{TP + FP} \quad (4)$$

Recall_2 indicates the proportion of all samples whose true labels are positive samples that are predicted to be positive samples. The value ranges from 0 to 1. The larger the value, the better the prediction effect:

$$\text{Recall}_2 = \frac{TP}{TP + FN} \quad (5)$$

F1 is the harmonic mean of Precision_2 and Recall_2 , ranging from 0 to 1. The larger the value, the better the prediction effect:

$$F1 = \frac{2 \times \text{Precision}_2 \times \text{Recall}_2}{\text{Precision}_2 + \text{Recall}_2} \quad (6)$$

IoU represents the ratio of the intersection and union between the prediction result of a certain category and the real label. The value ranges from 0 to 1. The larger the value, the better the prediction effect:

$$IoU = \frac{TP}{TP + FP + FN} \quad (7)$$

4. Experimental Process and Result Analysis

4.1. Parameter Setting of Scene Classification Experiment

The first layer of the hierarchical information extraction method is to use the deep learning scene classification EfficientNet-B5 model to locate the photovoltaic power plants on medium-resolution remote sensing images.

We implemented our network on the PyTorch [35] framework and executed it on a 64-bit Ubuntu 20.04 computer with 12 GB memory NVIDIA TITAN V. This experiment uses the EfficientNet-B5 model pre-trained on the tens of millions of ImageNet datasets [36]. The input and output layers of the model are modified to make the model suitable for eight-band image input and binary classification output. A total of 22,025 samples with a 128×128 pixel size are randomly divided into the train set, validation set and test set input model in a ratio of 7:1:2. In this work, we use the stochastic gradient descent (SGD) optimizer with an initial learning rate of 0.005, a momentum of 0.9, and a weight decay of 0.0001. The experiment adopts a fixed-step size-decay strategy [37] for the learning rate parameter; the learning rate is changed to 0.1 times the original every 30 epochs, and a total of 100 epochs are trained. During the model training process, the accuracy of the validation set reached a maximum of 99.97%, and the model was saved for prediction. A large-scale image is input into the model in the form of a fixed window sliding, and finally the areas containing the photovoltaic panels are obtained.

4.2. Scene Classification Results and Analysis

The spatial information positioning results of the photovoltaic power plants in Qinghai Province, Gansu Province and Xinjiang Uygur Autonomous Region are shown in Figure 8a, Figure 8b, and Figure 8c, respectively.

The red vector range is the provincial administrative boundary of each province, the blue vector range is the visually interpreted area containing photovoltaic panels and only represents a rough positioning, and the green vector range is the area predicted by the model or suspected to contain photovoltaic panels. The purpose of this step is to locate the spatial information of the photovoltaic panels and to reduce the false negatives of the photovoltaic panels on the basis of ensuring that the photovoltaic panels are not missed.

In Qinghai Province, a total of 18 photovoltaic power plants were input into the model as the training dataset, and the model detected a total of 22 photovoltaic power plants, of which the number of undetected photovoltaic power plants in the training dataset was 0. In Gansu Province, a total of 36 photovoltaic power plants were input into the model as the training dataset, and the model detected a total of 44 photovoltaic power plants, of which the number of undetected photovoltaic power plants in the training dataset was 0. Finally, in Xinjiang Uygur Autonomous Region, a total of 88 photovoltaic power plants were input into the model as the training dataset, and the model detected a total of 118 photovoltaic power plants, of which the number of undetected photovoltaic power plants in the training dataset was 0.

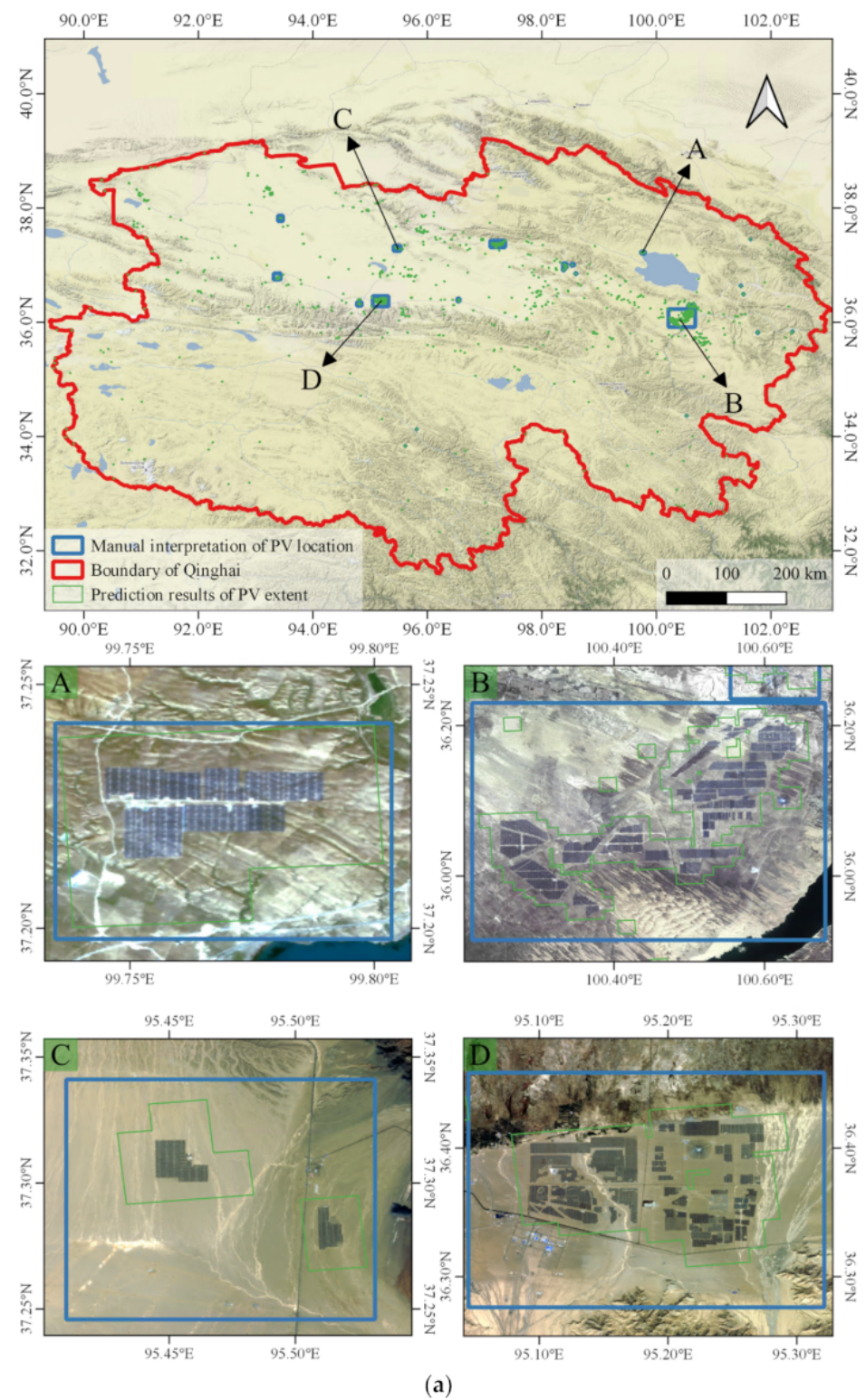


Figure 8. Cont.

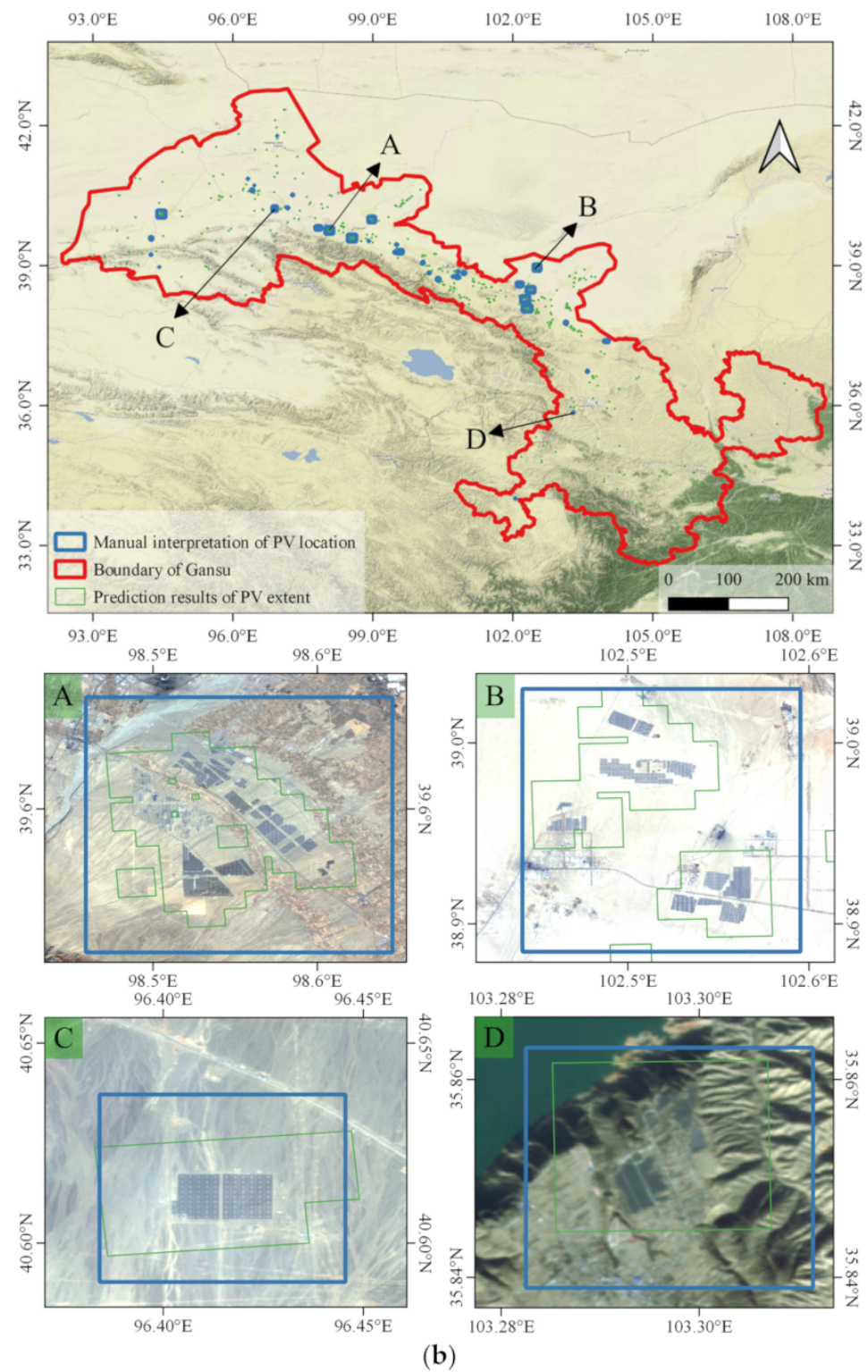


Figure 8. Cont.

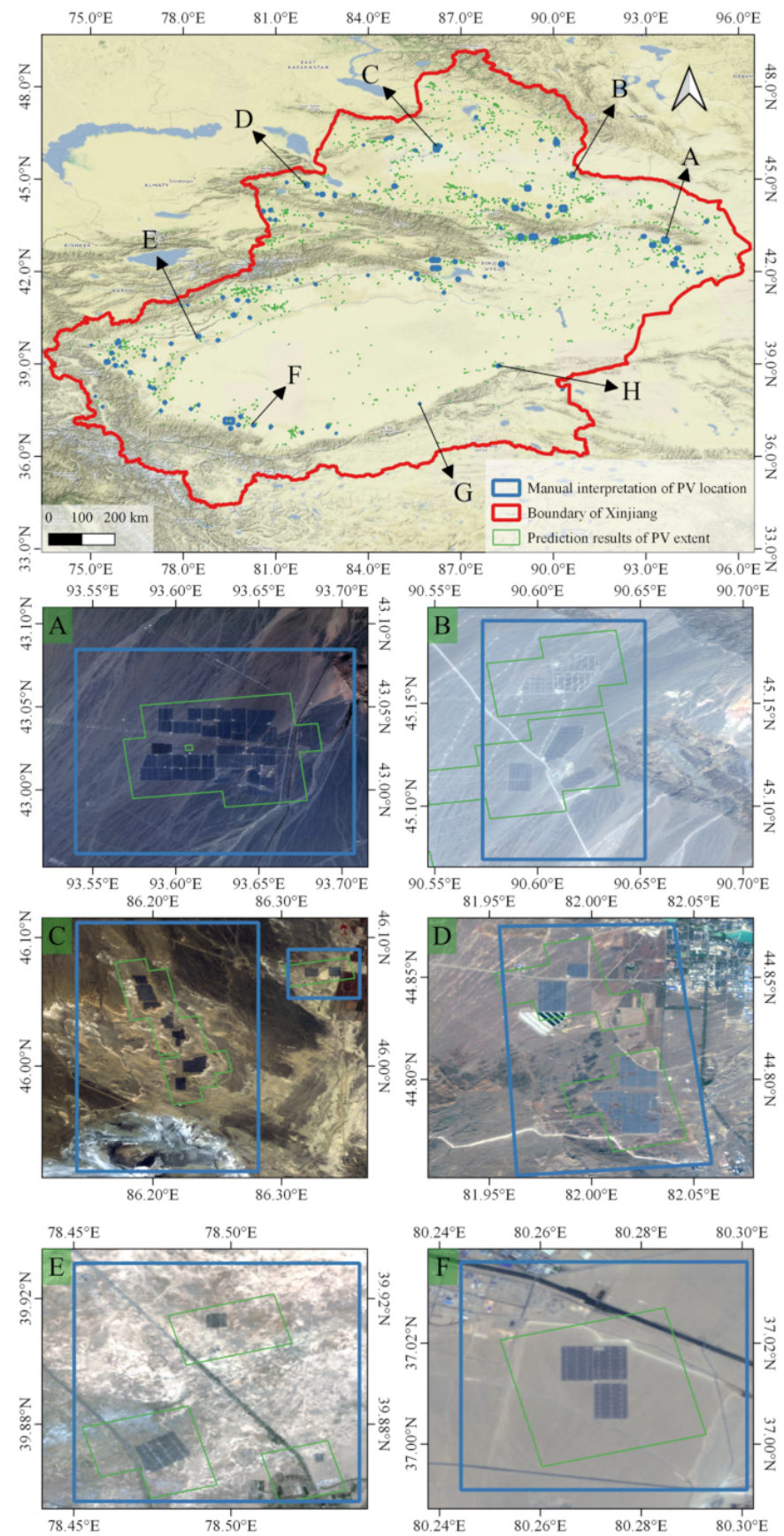


Figure 8. Cont.

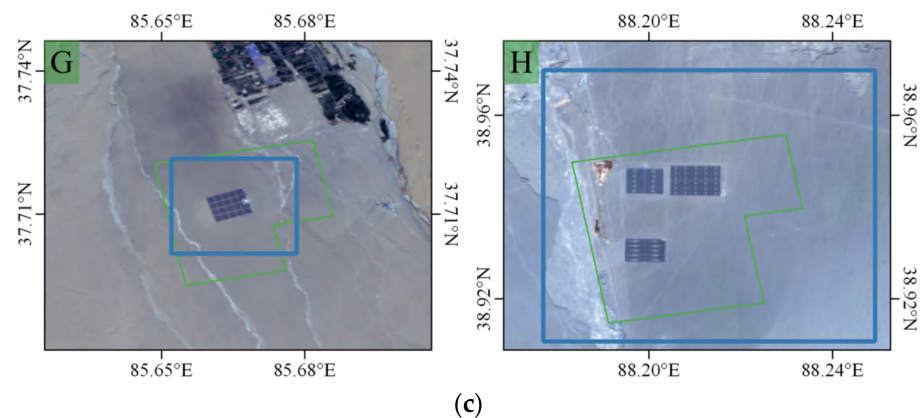


Figure 8. Spatial information positioning results of photovoltaic power plants in the study area. (a) Spatial information positioning results of photovoltaic power plants in Qinghai Province. (b) Spatial information positioning results of photovoltaic power plants in Gansu Province. (c) Spatial information positioning results of photovoltaic power plants in Xinjiang Uygur Autonomous Region.

Gansu Province was selected as the test area in this experiment, and the scene classification results in Gansu Province were visually interpreted. Through manual interpretation, there are 1029 scenes predicted by the model to include photovoltaics in Gansu Province. There are 549 scenes that actually contain photovoltaics in the prediction results and 480 scenes that do not contain photovoltaics. It is worth mentioning that the forecast results include all photovoltaics in Gansu Province.

$$\text{Precision}_1 = \frac{N_1}{N_1 + N_2} \times 100\% = \frac{549}{549 + 480} \times 100\% = 53.4\% \quad (8)$$

$$\text{Recall}_1 = \frac{N_1}{N_1 + N_3} \times 100\% = \frac{549}{549 + 0} \times 100\% = 100\% \quad (9)$$

The evaluation index Recall_1 is 100%, indicating that all photovoltaic power plants in Gansu Province have been detected, and there is no effect on the precise identification steps of photovoltaic panels; the evaluation index Precision_1 is 54.3%, indicating that under the premise of excluding non-PV scenes to a large extent, the accuracy rate of the prediction results is still more than half.

This paper demonstrates the feasibility of the method through experiments. This method largely eliminates areas without photovoltaic panels and largely preserves areas where photovoltaic panels are present. Therefore, this method can be used for practical promotion and application.

4.3. Parameter Setting of Semantic Segmentation Experiment

The second layer of the hierarchical information extraction method is to use the deep learning semantic segmentation U²-Net model to extract the photovoltaic panels precisely on high-resolution remote sensing images.

We implemented our network on the PyTorch [35] framework and executed it on a 64-bit Ubuntu 20.04 computer with a 12 GB-memory NVIDIA TITAN V. This experiment uses the double-nested U²-Net model. In total, 10,000 pairs of 512 × 512 pixel size samples with a spatial resolution of 2 m were randomly divided into a training set, validation set and test set input model in a ratio of 16:5:5. In this work, we use the stochastic gradient descent (SGD) optimizer with an initial learning rate of 0.001, a momentum of 0.9 and a weight decay of 0.0001. The experiment adopts a fixed-step-size decay strategy [37] for the learning rate parameter; the learning rate is changed to 0.5 times the original every 20 epochs, and a total of 100 epochs are trained. During model training, the highest accuracy on the validation set was 97.415%, and the model was saved for prediction. The small-scale two-meter spatial

resolution image including the photovoltaic panel area is input into the model, and the precise distribution of the photovoltaic panels is finally obtained.

4.4. Semantic Segmentation Results and Analysis

The precise extraction result of the photovoltaic panels in Qinghai Province, Gansu Province and Xinjiang Uygur Autonomous Region are shown in Figure 10a, Figure 10b, and Figure 10c, respectively:

The red vector range is the provincial administrative boundary of each province, and the red grid range is the result of the precise extraction of photovoltaic panels. The performance of the model on the test set is shown in Table 5. The extraction results of the centralized photovoltaic power plants in this experiment can provide data support for the estimation of photovoltaic power generation, the macro-control of the photovoltaic industry, and national decision-making.

Table 5. Performance of the model on the test set.

Accuracy	Precision ₂	Recall ₂	F1	IoU
97.686%	0.937	0.964	0.950	0.906

The overall accuracy of the model on the test set is 97.686%. This method realizes high-precision and rapid extraction of photovoltaic panels in a large area.

In the process of scene classification based on remote sensing images with a spatial resolution of 16 m, there are some confusion classes, including regularly shaped farmland, shaded and regularly textured bare ground, and regular shaped water. However, these confusion classes are almost eliminated in the second step, semantic segmentation, of the proposed method.

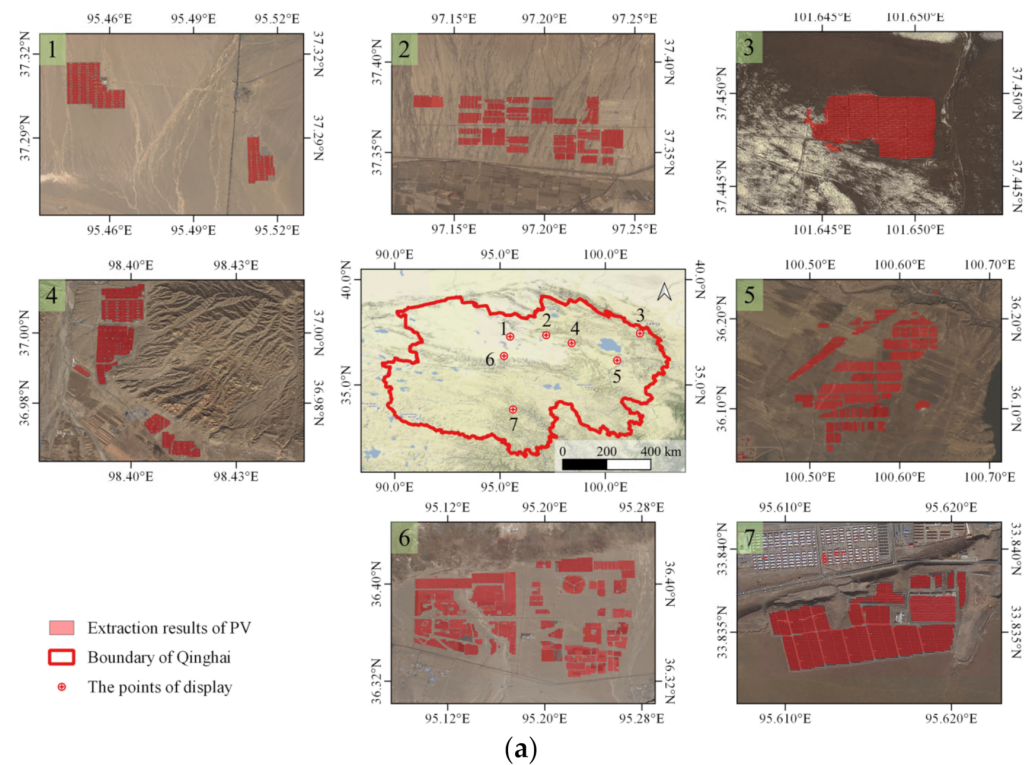


Figure 9. Cont.

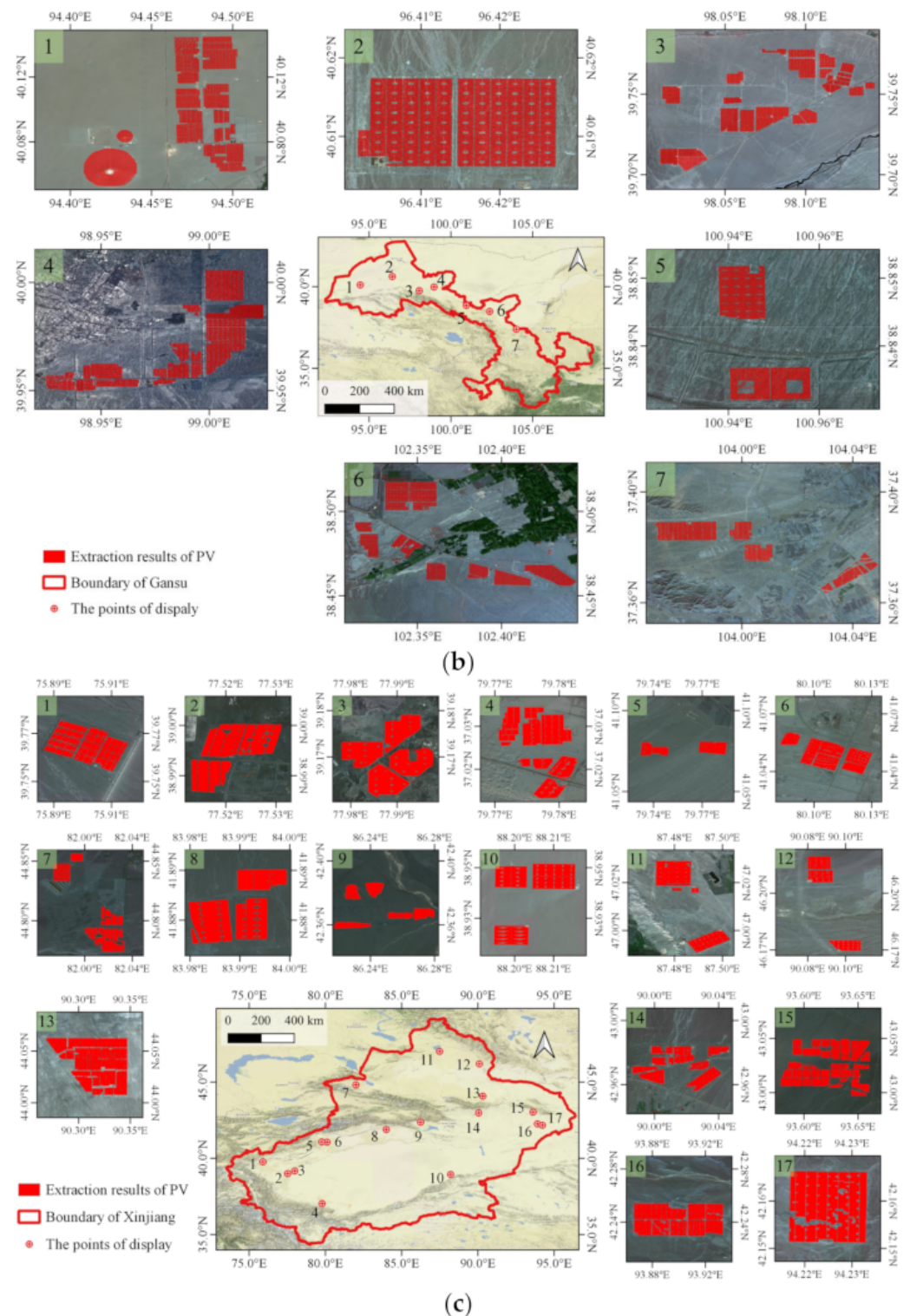


Figure 10. Precise extraction results in the study area. (a) Precise extraction results of the photovoltaic panels in Qinghai Province. (b) Precise extraction results of the photovoltaic panels in Gansu Province. (c) Precise extraction results of the photovoltaic panels in Xinjiang Uygur Autonomous Region.

Three main classes that are easily misclassified as scenes containing photovoltaic panels are shown in Figure 11. The first row of images are scenes that were mistaken for containing photovoltaic panels in the 16-m spatial resolution remote sensing image; the second row of images are the 2-m spatial resolution remote sensing images corresponding to the first row of images; the third row of images are extraction results.

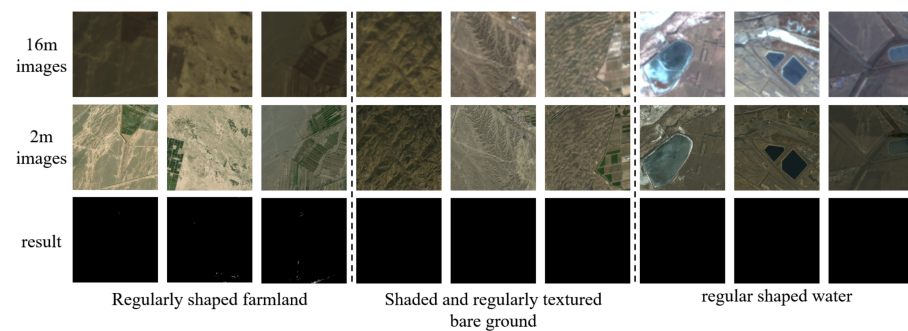


Figure 11. Scene classes that are easily confused during scene classification.

Additionally, Section 4.5 of this paper will verify the superiority of the method proposed in this paper in terms of accuracy and efficiency through two sets of comparative experiments.

4.5. Comparative Experiments

In order to prove the superiority of the method proposed in this paper, two sets of comparative experiments are designed.

First, in order to demonstrate the superiority of the method proposed in this paper in terms of accuracy, the method proposed in this paper is compared with four classical methods including DeepLabV3+, U-Net, SegNet, and FCN8s to conduct comparative experiments on the same dataset. The dataset is divided into a training set, validation set, and test set according to the ratio of 16:5:5. The quantitative results of the five methods on the test set are shown in Table 6.

Table 6. The quantitative results of different methods.

Method	Accuracy	Precision ₂	Recall ₂	F1	IoU
FCN8s	97.589%	0.952	0.946	0.949	0.903
SegNet	97.199%	0.915	0.965	0.939	0.886
U-Net	96.756%	0.900	0.961	0.929	0.868
DeepLabV3+	87.269%	0.926	0.667	0.776	0.633
Our method	97.686%	0.937	0.964	0.950	0.906

It can be seen from the data in Table 6 that the hierarchical information extraction method proposed in this paper has the highest accuracy. F1 is the index for comprehensive evaluation of Precision₂ and Recall₂, and the method proposed in this paper also has the highest F1. IoU is typically reported as the top-level performance of a semantic segmentation model, and the method proposed in this paper still has the highest IoU.

Therefore, the hierarchical information extraction method has superiority in accuracy in large-scale photovoltaic identification tasks.

Second, in order to demonstrate the superiority of the method proposed in this paper in terms of efficiency, Gansu Province was selected as the test area for comparative experiments. Five methods, including the method proposed in this paper, FCN8s, SegNet, U-Net, and DeepLabV3+, were used to identify photovoltaic panels in Gansu Province. The time taken by each method is shown in Table 7.

Table 7. Time spent for PV panel extraction on the test area by different methods.

Method	Our Method	FCN8s	SegNet	U-Net	DeepLabV3+
Time	7.317 h	12.333 h	12.917 h	12.767 h	28.733 h

It can be seen from Table 7 that the method proposed in this paper takes the shortest time of 7.317 h and has the highest efficiency. The shortest time of the other four methods is 12.333 h

for FCN8s; the longest time of the other four methods is 28.733 h for DeepLabV3+. Compared with the FCN8s, the method proposed in this paper improves the efficiency by 40.7%.

Therefore, the hierarchical information extraction method has superiority in time in large-scale photovoltaic identification tasks.

5. Conclusions

This study proposed a hierarchical information extraction method for large-scale centralized photovoltaic power plants based on multi-source remote sensing images. This method takes full advantage of the characteristics of remote sensing images with different spatial resolutions. Firstly, the spatial information positioning of photovoltaic panels is quickly realized in the remote sensing image with a spatial resolution of 16 m. The advantages of this step are as follows: (1) The area that does not contain photovoltaic power plants is greatly reduced, the number of positive and negative targets is balanced, and the impact of the imbalance of positive and negative samples on the precise extraction of photovoltaic panels is reduced. (2) The input range of high-resolution remote sensing images is greatly reduced, the waste of resources is avoided, and the efficiency and accuracy of the identification of photovoltaic panels are improved. Secondly, the accurate identification of photovoltaic panels is realized in remote sensing images with a spatial resolution of 2 m. This step can assist the refined management of large photovoltaic power plants and provide data and information support for the country to make relevant decisions and global sustainable development.

This study realizes the distribution mapping of large-scale and high-resolution centralized photovoltaic power panels based on experiments in three northwestern Chinese provinces.

In summary, clean energy such as solar energy plays an important role in the era of sustainable development. Accurately obtaining the location, spatial distribution, and area information of photovoltaic power plants is of great significance for optimizing the energy structure and rationally exploiting non-renewable energy. Remote sensing provides a new way to objectively and impartially obtain the production capacity of photovoltaic power plants through non-contact, long-distance, and large-scale measurements. The hierarchical information extraction method for centralized photovoltaic power plants based on multi-source remote sensing images is effective in large-scale range. Finally, it will definitely provide technical support for related industries and can be popularized and applied.

Author Contributions: Methodology, F.G., G.H. and G.W.; datasets, F.G., G.W. and D.Z.; experiments, F.G., D.Z. and L.T.; mapping the experimental results, F.G., G.H., G.W., R.Y. and D.Z.; results analysis, F.G., G.H. and R.Y.; data curation, F.G.; writing—original draft preparation, F.G.; writing—review and editing, F.G., G.H. and D.Z.; project administration, G.H. and G.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by a subject of the Second Tibetan Plateau Scientific Expedition and Research Program (STEP), grant number 2019QZKK030701; the Strategic Priority Research Program of the Chinese Academy of Sciences, grant number XDA19090300; and the National Natural Science Foundation of China, grant number 61731022.

Data Availability Statement: Not applicable.

Acknowledgments: The authors thank the anonymous reviewers and the editors for their valuable comments to improve our manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Guo, S. Development Status and Prospect of Solar Photovoltaic Power Generation. *Shandong Ind. Technol.* **2018**, *16*, 163.
2. Cao, S.; Zhou, G.; Cai, Q.; Ye, Q.; Pang, H. Solar Cell Review: Materials, Policy-Driven Mechanisms and Application Prospects. *Acta Mater. Compos. Sin.* **2022**, *39*, 1847–1858.
3. Li, Y.; Zhang, Q.; Wang, G.; McLellan, B.; Liu, X.; Wang, L. A review of photovoltaic poverty alleviation projects in China: Current status, challenge and policy recommendations. *Renew. Sust. Energ. Rev.* **2018**, *94*, 214–223. [[CrossRef](#)]

4. Zhang, H.; Xu, Z.; Sun, C.; Elahi, E. Targeted poverty alleviation using photovoltaic power: Review of Chinese policies. *Energy Policy* **2018**, *120*, 550–558. [CrossRef]
5. Padmanathan, K.; Govindarajan, U.; Ramachandaramurthy, V.K.; Rajagopalan, A.; Pachaivannan, N.; Sowmmiya, U.; Periasamy, S.K. A sociocultural study on solar photovoltaic energy system in India: Stratification and policy implication. *J. Clean. Prod.* **2019**, *216*, 461–481. [CrossRef]
6. Coria, G.; Penizzotto, F.; Pringles, R. Economic analysis of photovoltaic projects: The Argentinian renewable generation policy for residential sectors. *Renew. Energy* **2019**, *133*, 1167–1177. [CrossRef]
7. Jan, I.; Ullah, W.; Ashfaq, M. Social acceptability of solar photovoltaic system in Pakistan: Key determinants and policy implications. *J. Clean. Prod.* **2020**, *274*, 123140. [CrossRef]
8. International Energy Agency. Available online: <https://www.iea.org/reports/renewables-2020/solar-pv> (accessed on 5 June 2022).
9. Central People's Government of the People's Republic of China. Available online: http://www.gov.cn/xinwen/2022-01/22/content_5669854.htm (accessed on 5 June 2022).
10. Wang, S.; Zhu, S.; Jiang, Y. OLI image photovoltaic panel scene recognition based on CNN model transfer. *Bull. Surv. Mapp.* **2022**, *2*, 5–9.
11. Wang, S. Application of machine learning method in remote sensing extraction of photovoltaic power station. Master's Thesis, Jiangsu Normal University, Xuzhou, China, 24 May 2018.
12. Liang, G. Research on Segmentation Methods of Small Objects in Images. Master's Thesis, University of Electronic Science and Technology of China, Chengdu, China, 30 May 2020.
13. Zhang, X.; Zeraatpisheh, M.; Rahman, M.M.; Wang, S.; Xu, M. Texture is important in improving the accuracy of mapping photovoltaic power plants: A case study of Ningxia Autonomous Region, China. *Remote Sens.* **2021**, *13*, 3909. [CrossRef]
14. Li, Y.; Liu, Y. Photovoltaic field extraction of multi-source remote sensing images supported by object image analysis method. *Geomat. Spat. Inf. Technol.* **2016**, *3*, 68–72.
15. Wang, S.; Zhang, L.; Li, H.; Zhao, Z.; Shen, Y.; Chai, Q. Research on improved optimal band combination method for specific target feature extraction. *Bull. Surv. Mapp.* **2017**, *7*, 49–54.
16. Costa, M.V.C.V.D.; Carvalho, O.L.F.D.; Orlandi, A.G.; Hirata, I.; Albuquerque, A.O.D.; Silva, F.V.E.; Júnior, O.A.D.C. Remote sensing for monitoring photovoltaic solar plants in Brazil using deep semantic segmentation. *Energies* **2021**, *14*, 2960. [CrossRef]
17. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015.
18. Chen, L.C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European Conference on Computer Vision, Munich, Germany, 8–14 September 2018.
19. Lin, T.Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature pyramid networks for object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition, Honolulu, HI, USA, 21–26 July 2017.
20. Zhao, H.; Shi, J.; Qi, X.; Wang, X.; Jia, J. Pyramid scene parsing network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
21. Zhang, X.; Zhou, Y.N.; Luo, J. Deep learning for processing and analysis of remote sensing big data: A technical review. *Big Earth Data* **2021**, 1–34. [CrossRef]
22. Zhou, D.; Wang, G.; He, G.; Yin, R.; Long, T.; Zhang, Z.; Chen, S.; Luo, B. A Large-Scale Mapping Scheme for Urban Building From Gaofen-2 Images Using Deep Learning and Hierarchical Approach. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 11530–11545. [CrossRef]
23. Central People's Government of the People's Republic of China. Available online: http://www.gov.cn/test/2005-06/15/content_18253.htm (accessed on 6 June 2022).
24. National Energy Administration. Available online: http://www.nea.gov.cn/2022-03/09/c_1310508114.htm (accessed on 6 June 2022).
25. Huang, Z.; Fang, H.; Li, Q.; Li, Z.; Zhang, T.; Sang, N.; Li, Y. Optical remote sensing image enhancement with weak structure preservation via spatially adaptive gamma correction. *Infrared Phys. Technol.* **2018**, *94*, 38–47. [CrossRef]
26. Bhandari, A.K. A logarithmic law based histogram modification scheme for naturalness image contrast enhancement. *J. Ambient Intell. Humaniz. Comput.* **2020**, *11*, 1605–1627. [CrossRef]
27. Nnolim, U.A. An adaptive RGB colour enhancement formulation for logarithmic image processing-based algorithms. *Optik* **2018**, *154*, 192–215. [CrossRef]
28. Tan, M.; Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In Proceedings of the International Conference on Machine Learning, Long Beach, CA, USA, 9–15 June 2019.
29. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.C. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018.
30. Cheng, D.; Meng, G.; Cheng, G.; Pan, C. SeNet: Structured edge network for sea-land segmentation. *IEEE Geosci. Remote. Sens. Lett.* **2016**, *14*, 247–251. [CrossRef]
31. Targ, S.; Almeida, D.; Lyman, K. Resnet in resnet: Generalizing residual architectures. *arXiv* **2016**, arXiv:1603.08029.
32. Iandola, F.; Moskewicz, M.; Karayev, S.; Girshick, R.; Darrell, T.; Keutzer, K. Densenet: Implementing efficient convnet descriptor pyramids. *arXiv* **2014**, arXiv:1404.1869.
33. Qin, X.; Zhang, Z.; Huang, C.; Dehghan, M.; Zaiane, O.R.; Jagersand, M. U²-Net: Going deeper with nested U-structure for salient object detection. *Pattern Recognit.* **2020**, *106*, 107404. [CrossRef]

-
34. Luque, A.; Carrasco, A.; Martín, A.; de Las Heras, A. The impact of class imbalance in classification performance metrics based on the binary confusion matrix. *Pattern Recognit.* **2019**, *91*, 216–231. [[CrossRef](#)]
 35. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Chintala, S. Pytorch: An imperative style, high-performance deep learning library. In Proceedings of the Advances in neural information processing systems, Vancouver, BC, Canada, 8–14 December 2019.
 36. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009.
 37. You, K.; Long, M.; Wang, J.; Jordan, M.I. How does learning rate decay help modern neural networks? *Learning* **2019**, *10*, 2–10.